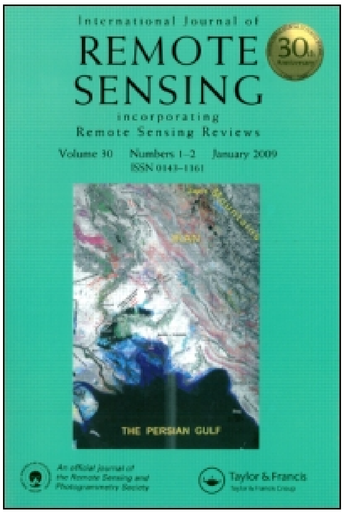

International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713722504

Modelling and prediction of malaria vector distribution in Bangladesh from remote-sensing data

A. Rahman[a]; F. Kogan[b]; L. Roytman[a]; M. Goldberg[b]; W. Guo[c]
[a] NOAA-CREST, the City College of New York, New York, NY, USA [b] NOAA/NESDIS, MD, USA [c] IMSG, MD, US

Online publication date: 22 March 2011

To cite this Article Rahman, A. , Kogan, F. , Roytman, L. , Goldberg, M. and Guo, W.(2011) 'Modelling and prediction of malaria vector distribution in Bangladesh from remote-sensing data', International Journal of Remote Sensing, 32: 5, 1233 — 1251

To link to this Article: DOI: 10.1080/01431160903527447
URL: http://dx.doi.org/10.1080/01431160903527447

International Journal of Remote Sensing
Vol. 32, No. 5, 10 March 2011, 1233–1251

Taylor & Francis
Taylor & Francis Group

Modelling and prediction of malaria vector distribution in Bangladesh from remote-sensing data

A. RAHMAN*†, F. KOGAN‡, L. ROYTMAN†, M. GOLDBERG‡ and W. GUO§
†NOAA-CREST, the City College of New York, New York, NY 10031, USA
‡NOAA/NESDIS, 5200 Auth Road Camp Spring, MD 20746, USA
§IMSG, 5200 Auth Road Camp Spring, MD 20746, US

(*Received 7 December 2008; in final form 25 October 2009*)

Epidemic malaria cases and satellite-based vegetation health (VH) indices were investigated to be used as predictors of malaria vector activities in Bangladesh. The VH indices were derived from radiances, measured by the Advanced Very High Resolution Radiometer (AVHRR) on National Oceanic and Atmospheric Administration (NOAA) afternoon polar orbiting satellites. Two indices characterizing moisture and thermal conditions were investigated using correlation and regression analysis applied to the number of malaria cases recorded in the entire Bangladesh region and three administrative divisions (Chittagong, Sylhet and Dhaka) during 1992–2001. It is shown that during the cooler months (November to March), when mosquitoes are less active, the correlation between number of malaria cases and two investigated indices was near zero. From April, when the mosquito activity season starts, the correlation increased, reaching a maximum value of 0.5–0.8 by the middle of the high season (June to July), reducing thereafter to zero by the beginning of the cool season in November. Following these results, regressional equations for the number of malaria cases as a function of VH indices were built and tested independently. They showed that, in the main malaria administrative division (Chittagong) and the entire Bangladesh region, the regressional equations can be used for early prediction of malaria development.

1. Introduction

Malaria is endemic to over 100 countries around the world, and is responsible for over 300 to 500 million clinical cases and more than a million deaths each year (Montanari *et al.* 2001, Faiz *et al.* 2002). Nearly 40% of the world's population, living mostly in the poorest countries, is at risk of malaria each year.

Malaria has been known to cause febrile illness in Bangladesh for a long time. Nearly 200 000 malaria cases are reported each year in Bangladesh for a population of 140 million. This number can fluctuate depending on weather conditions (Elias and Rahman 1987, Githeko *et al.* 2000, Craig *et al.* 2004). Malaria transmission in Bangladesh is mostly seasonal and limited to the border regions with Myanmar in the east and India in the north (http://www.bangladeshgov.org/bdmaps) (see figure 1). Out of the country's six administrative divisions (containing 64 districts), Dhaka, Sylhet and Chittagong (13 districts) are malaria endemic (Najera *et al.* 1998, WHO 2002, Ingrid and Van 2004). These three divisions contribute nearly 98% of the total

*Corresponding author. Email: fmatiq@yahoo.com

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160903527447

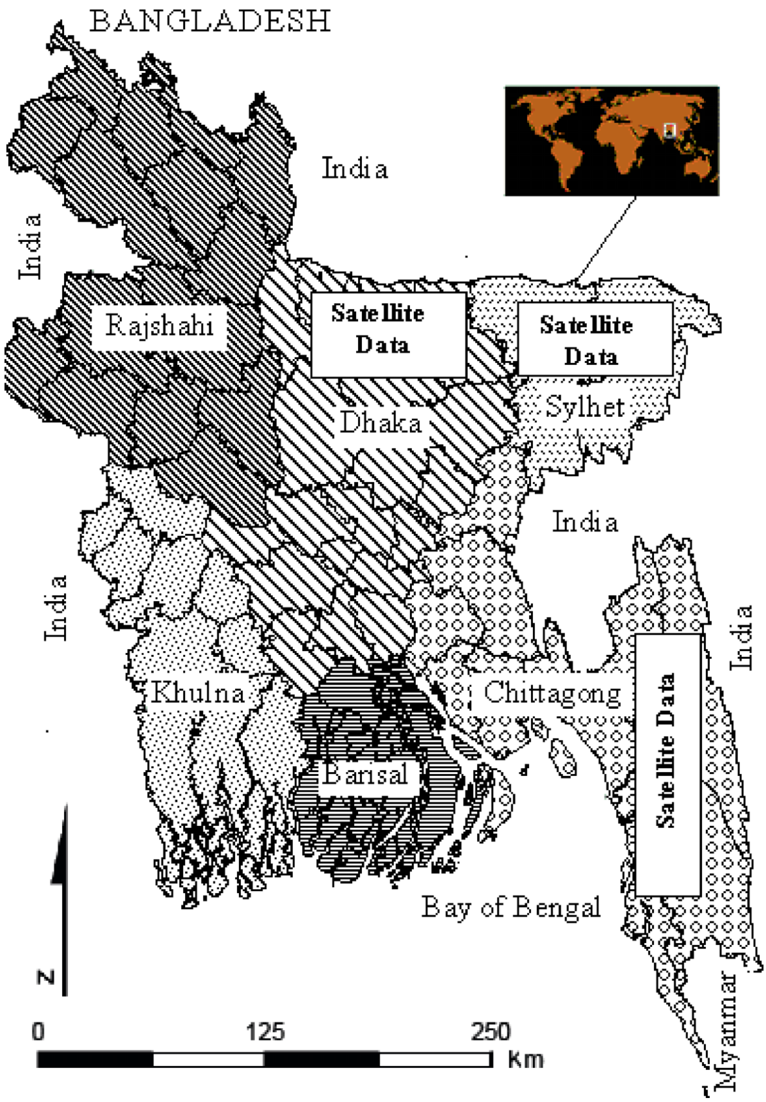

Figure 1. Administrative divisions and malaria endemic area where satellite data were collected in Bangladesh.

Bangladesh malaria morbidity and mortality statistics reported each year (Rosenberg *et al.* 1982). Around 27 million people (20% of the total Bangladesh population) live in a malaria endemic area (Gramiccia 1952, Rahman *et al.* 2006). Intense perennial malaria transmission peaks during the rainy season.

The mosquito's development, activity and ability to transmit malaria changes from year to year, depending on the weather conditions. Therefore, weather parameters are often used as indicators to monitor the malaria epidemic (Nagpal and Sharma 1995, Smith and McKenzie 2004, Zhou *et al.* 2004). There are 34 weather stations for the 144000 km^2 area of Bangladesh; each station covers nearly 4000 km^2. This area is too large for efficient malaria monitoring. Moreover, the stations are not equally distributed, with a particularly low density in malaria prone areas. Therefore, information from weather stations is not sufficient for effective malaria monitoring in Bangladesh. This paper investigates potential for application of 4 km Advanced Very High Resolution Radiometer (AVHRR) data obtained from National Oceanic and Atmospheric Administration (NOAA) operational environmental satellites (Kogan 2001) for this purpose. Specifically, vegetation health (VH) indices were applied for detection, surveillance and numerical estimation of malaria development (Byron *et al.* 1991, Connor *et al*. 1999, Kaya *et al*. 2002).

2. Environment

Bangladesh is mostly low, flat alluvial plains (with hills in the southeast) intersected by numerous rivers, rivulets, canals, swamps and marshy lands. Arable land accounts for 61%, forest for 8%, water for 9% and other landscapes for 22% of the total Bangladesh area.

The climate of Bangladesh is sub-tropical warm, wet and humid (Pampana 1969, Rahman *et al.* 2006). In the malaria endemic areas, the annual rainfall varies between 1000 and 3000 mm, and temperature and humidity ranges are 12–30°C and 65–90% (Rahman *et al.* 2006, Paresul 2008), respectively. There are two malaria seasons in Bangladesh: high from April through to October when the weather is wet and warm, and low from November through to March when the weather is cool and dry (Remme *et al.* 2001, Paresul 2008).

3. Mosquitoes, malaria and climate

Malaria parasite Plasmodium in Bangladesh is transmitted by female Anopheles mosquitoes (Rosenberg and Maheswary 1982, Elias and Rahman 1987, Ingrid and Van 2004). Out of the 15 mosquito species, Anopheles Dirus (AD) is the most widely spread in Southeast Asia, including Bangladesh. Mosquitoes in Bangladesh transmit malaria year-round (Gramiccia 1952). However, during the cooler season (November through to March), mosquitoes are less active and the number of malaria cases is relatively small. This number increases considerably during the warm and wet season (April through to October) (Rahman *et al.* 2006, Paresul 2008).

Malaria is transmitted by infected adult female mosquitoes that bite to get blood for laying eggs. The mosquito hatching period from laying eggs to an adult stage is from 7–15 days. An entire cycle, when the AD is able to bite, transmit the parasite and malaria can be observed is 15–50 days (Pampana 1969, Boëte and Koella 2002). Therefore, during April to October, four to five cycles of mosquito population are able to transmit malaria. The incubation period for development of malaria after infected mosquito bites is between 8 and 35 days.

Most research points out three climate factors controlling mosquito activity and their ability to transmit malaria: rainfall, temperature and humidity (Pampana 1969). The optimum temperatures for malaria development and activity are 25–27°C (Hay *et al.* 2002, Bouma 2003). If the daytime temperature exceeds 40°C, mosquitoes are less active and parasite transmission is very limited. In general, a larger amount of rainfall stimulates their activity. However, frequent and intensive rainfall during the monsoon period might produce stagnation in malaria transmission since it washes out eggs and reduces the chances for development of adult mosquitoes (Thomson and Connor 2001, Chandramohan *et al.* 2002). In the environment of Bangladesh, AD females stay active during the period when precipitation exceeds 50 mm per month.

However, a combination of large rainfall and hot weather during June to August might reduce mosquito activity (Chilundo *et al.* 2004, Rahman *et al.* 2006). Also, malaria transmission might slow down if the humidity drops below 60%.

In a number of studies, weather parameters (rainfall, temperature and humidity) were used to monitor and predict mosquito development, activities and malaria transmission (McMichael *et al.* 1996, Thomson *et al.* 2000, Rogers *et al.* 2002). There are also a few studies showing that the number of malaria cases correlated with satellite-measured parameters characterizing vegetation (Hay *et al.* 2001, Ceccato *et al.* 2005,). In our earlier research, it has been also shown that the number of malaria cases in the entire Bangladesh region correlates with AVHRR-based VH indices (Rahman *et al.* 2006). In this paper, we present regional analysis in application of VH indices for early prediction and monitoring of malaria epidemics.

4. Data

Regional malaria statistics and satellite data were used in this study. Malaria statistics were presented using the annual number of clinical malaria cases during 1992–2001. These data, collected from all Bangladesh hospitals, were obtained from the Directorate General of Health, Bangladesh's Ministry of Health. The hospital data were aggregated to local administrative unit health centres and further on to the administrative districts, and finally to the administrative divisions and the entire Bangladesh region (Wickramasinghe *et al.* 2002, Paresul 2008). These statistics were presented by the number of persons (person's total; PT) who came to a hospital with fever and the number of positive malaria cases (PMC). The dynamics of the PT and

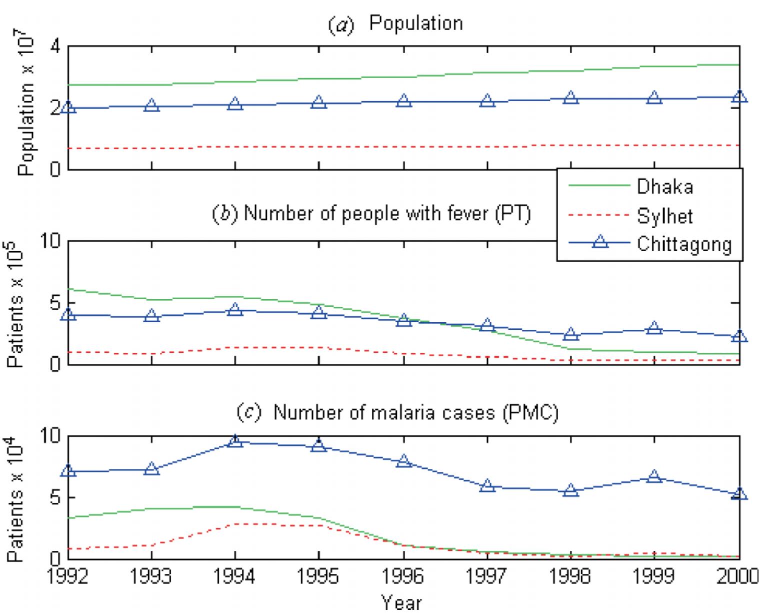

Figure 2. Yearly population increase, number of persons who come to hospital with fever and number of malaria cases for Dhaka, Sylhet and Chittagong divisions.

PMC during the investigated period, along with the total population, is shown in figure 2. Although the population in regions does not change considerably, the PT and PMC slightly decreases, indicating that the Government took some measures to improve the health of the people. However, if the PMC is expressed as a percentage of the PT of what is normally done in malaria research, then the malaria dynamics show an increasing trend for the entire Bangladesh region and the major region (Chittagong) and a decreasing trend for the minor malaria regions.

Satellite data were collected from the NOAA/National Environmental Satellite, Data and Information Service (NESDIS) global vegetation index (GVI) dataset during 1992–2001. The GVI has a spatial resolution of 4×4 km, sampled to 16 km^2 and a daily temporal resolution sampled to a 7 day composite (Kidwell 1997). The GVI data contains reflectance in the visible (channel 1; ch1; 0.58–0.68 μm), near infrared (channel 2; ch2; 0.72–1.1 μm) and infrared (channel 4; ch4; 10.3–11.3 μm; and channel 5; ch5; 11.3–12.3 μm) spectral bands. Pre-launch and post-launch calibration coefficients were applied to ch1 and ch2 data in order to calculate reflectance. The normalized difference vegetation index (NDVI) was calculated as NDVI $= (R_2 - R_1) / (R_2 + R_1)$, where R_1 and R_2 represent measured radiances in the ch1 and ch2 bands, respectively.

The ch4 radiances were converted to brightness temperature (T_B), which was adjusted for non-linear behaviour of the AVHRR instrument. Satellite data were collected for Dhaka, Chittagong and Sylhet administrative divisions. In each of these divisions, the spatial average values of the NDVI and T_B were calculated for each week during 1992–2001 inside the area shown in figure 1. Satellite data for the entire Bangladesh region were aggregated as weighted averages for the three administrative divisions. The yearly coefficients for each administrative division were estimated from

$$C_i = E_i / B_i, \tag{1}$$

where C_i is the divisional average weighted coefficient in year i; E_i is the percentage of malaria cases for the division in year i; and B_i is the percentage of malaria cases for the entire Bangladesh region in year i.

Table 1 shows the annual coefficients used for each division to calculate the total Bangladesh percentage of malaria cases.

Table 1. Coefficients used to calculate the weighted mean of the VH indices for the entire Bangladesh region from the three administrative divisions, in percent (%).

	Administrative division		
Year	Dhaka	Chittagong	Sylhet
1992	0.30	0.63	0.07
1993	0.33	0.58	0.09
1994	0.26	0.57	0.17
1995	0.22	0.60	0.18
1996	0.11	0.79	0.10
1997	0.08	0.86	0.06
1998	0.05	0.91	0.04
1999	0.04	0.90	0.06
2000	0.03	0.93	0.04
2001	0.03	0.93	0.04

5. VH indices

The VH indices were developed from the NDVI and T_B. The data processing included removal of high-frequency noise from the annual time series of the NDVI and T_B, approximation of the annual cycle, calculation of the multi-year climatology and derivation of VH indices (Kogan 2001).

High-frequency temporal noise in the NDVI and T_B, related to the fluctuating transmission of the atmosphere, sun/sensor geometry, bi-directional reflectance, random noise and others, was removed by statistical smoothing of the NDVI and T_B annual time series for each pixel during the entire period using a combination of a median filter and the least-squares technique. The climatology of the NDVI and T_B seasonal cycle was approximated using the multi-year maximum (max) and minimum (min) weekly values taken from the smoothed data. The max and min for each pixel and week were calculated from 20 years of historical GVI data (Kogan 2002). The (max–min) criterion was used to describe and classify the weather-related ecosystem's 'carrying capacity' and therefore represented the climatology of those extreme weather-related fluctuations in the NDVI and T_B. The NDVI and T_B dynamics for the two years with the highest and lowest malaria cases for the three divisions of Bangladesh are shown in figure 3. The weekly NDVI and T_B for each year, together with the climatology, were used to approximate the VH indices, which are represented by the vegetation condition index (VCI) and the temperature condition index (TCI). Equations (2) and (3) show the numerical approximations of VCI and TCI values:

$$\mathrm{VCI} = 100[(\mathrm{NDVI}) - (\mathrm{NDVI}_{\min})]/[(\mathrm{NDVI}_{\max}) - (\mathrm{NDVI}_{\min})] \quad (2)$$

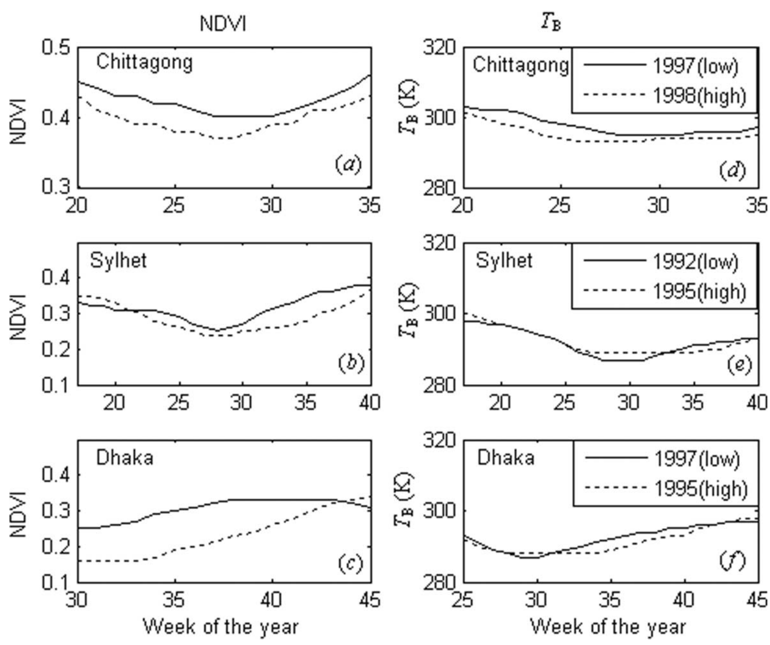

Figure 3. NDVI and T_B dynamics for the two years with highest and lowest malaria cases for three divisions of Bangladesh.

and

$$\mathrm{TCI} = 100[(T_{\mathrm{B\_max}}) - (T_{\mathrm{B}})]/[(T_{\mathrm{B\_max}}) - (T_{\mathrm{B\_min}})], \tag{3}$$

where NDVI, $\mathrm{NDVI}_{\mathrm{max}}$ and $\mathrm{NDVI}_{\mathrm{min}}$ (T_{B}, $T_{\mathrm{B\_max}}$ and $T_{\mathrm{B\_min}}$) are smoothed weekly NDVI (T_{B}) values and their multi-year absolute maximum and minimum values, respectively (Kogan 2001). The VCI and TCI values change from 0 to 100, reflecting changes in moisture and thermal conditions from extremely unfavourable (vegetation stress) to optimal (favourable). VCI and TCI values around 50 represent near-normal moisture (VCI) and thermal (TCI) conditions; VCI and TCI values below 50 indicate different levels of vegetation stress with the highest intensity equalling 0. On the opposite side of the scale, indices greater than 50 indicate no stress or favourable vegetation conditions (Rahman *et al.* 2006, Salazar *et al.* 2007).

6. Result and discussion

6.1 *Malaria dynamics*

Figure 4 shows the annual percentages of malaria cases and long-term trends in the three administrative divisions and the entire Bangladesh region during 1992–2001. Analysis of malaria dynamics indicates that the number of malaria cases experiences a long-term trend that is approximated by equation (4). Variations in the number of malaria cases around the trend are associated with weather fluctuations from year to year. They are expressed as a percentage deviation from the trend line (equation (5)) (Allard 1998, Rahman *et al.* 2006, Salazar *et al.* 2008):

$$Y_{\mathrm{trend}} = a_0 + a_1 T \tag{4}$$

and

$$D_i = (Y_i / Y_{\mathrm{trend}})100, \tag{5}$$

where Y_i are the observed malaria cases (% from the total number of people who come to local hospitals with fever) in year i; T is the year number; Y_{trend} is the number of malaria cases in a given year that fits the long-term trend straight line in a region

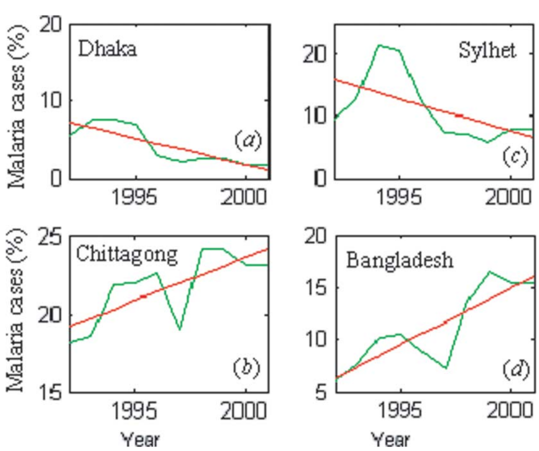

Figure 4. Annual malaria cases (expressed as % from the total number of people who come to the hospitals with fever) and trend line, 1992–2001.

Table 2. Slope and intercept for long-term trend approximation of malaria cases in the three administrative divisions.

Region	Intercept (a_0)	Slope (a_1)
Dhaka	7.94	–0.70
Chittagong	18.65	0.55
Sylhet	16.95	–1.03
Bangladesh	5.15	1.09

during 1992–2001; a_0 is an intercept; a_1 is a slope; D_i is a deviation from the trend (%) in year *i*. The parameters of the linear equations are shown in table 2.

Analysis of trends (as shown in figure 4) indicates that the total number of malaria cases in the entire Bangladesh region from 1992 to 2001 increased by 2.5 times (from 6 to 16%). This increase was associated with a similar trend in Chittagong division, which contributed 60% in the early 1990s to 93% in the early 2000s (table 1). The number of malaria cases in the other two malaria-prone Bangladesh divisions during the investigated period decreased. From 1992 to 2001, the Government of Bangladesh applied a few measures to eradicate malaria in the country. These measures resulted in a reduction of the number of malaria cases in Dhaka and Sylhet divisions. However, in the coastal zone of Chittagong division, with a disproportionally high poverty rate, the number of malaria cases expressed as percentages of people who come to hospital with fever kept increasing. That large weight of the Chittagong division in the malaria country total is mainly due to several factors: its geographic location in the coastal zone, the very high poverty rate and extremely poor sanitary conditions in many slum areas.

On an annual basis, the number of malaria cases fluctuates considerably. For example, in Chittagong division in 1997, the number of cases was smaller (the deviation, $D = 86\%$ or 14% below the trend); while in 1998, this number was larger ($D = 108\%$ or 8% above the trend). Even in divisions with a declining long-term trend, the *D* variation changes from year to year. For Sylhet division, the number of malaria cases in 1992 was 40% below the trend, while in 1995, it was 60% above the trend. Similarly in Dhaka division, the 1997 *D* was 46% below the trend, while in 1995, it was 34% above the trend. The *D* value for the entire Bangladesh region in 1997 was characterized by a small number of malaria cases ($D = 73\%$ or 27% below the trend), while in 1994 it was characterized by a large number of cases ($D = 116\%$ or 16% above the trend). Therefore, the assumption for the modelling was the following: years with *D* below the trend were unfavourable for mosquito development (less malaria transmission) and years with *D* above the trend were favourable (more malaria transmission). The next step includes correlation analysis of *D* with VH indices (Rahman *et al.* 2006).

6.2 *Correlation analysis*

Figure 5 shows the dynamics of the Pearson correlation coefficients (PCCs) between the end of each year *D* with weekly VCI and TCI values during 1992–2001. Analysis of the PCC in figure 5 indicates that there are two types of dynamics in the investigated areas. Dhaka and Sylhet divisions (contribute 3–4% of total Bangladesh malaria cases in the early 2000s, table 1) have erratic correlation dynamics; Chittagong division and the entire Bangladesh region have well-pronounced dynamics, corresponding to the

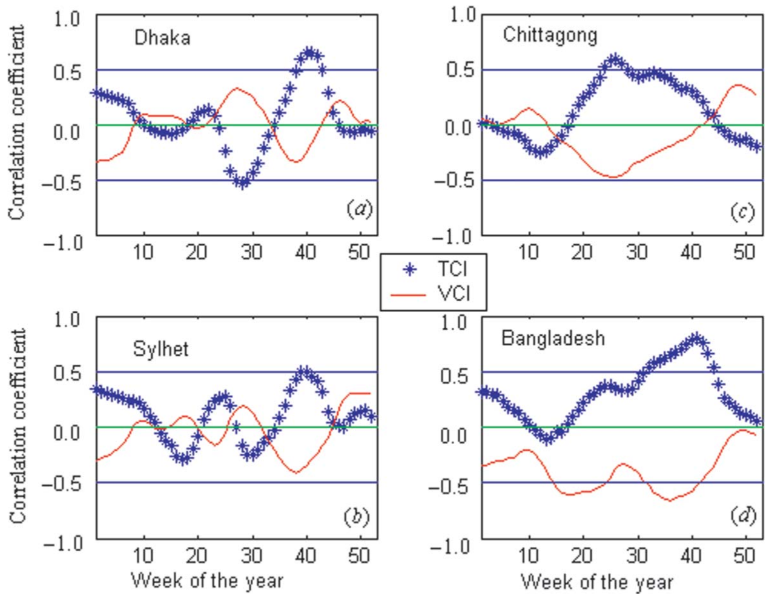

Figure 5. Pearson correlation coefficient dynamics of annual DY (percentage deviation of malaria cases from trend) versus weekly area-mean TCI and VCI.

main features of mosquito response to weather and correspondingly their ability to spread malaria. Following figure 5, during the cool season (November through to March) when the number of malaria cases is smaller, correlation of *D* with VCI and TCI values is low, indicating that VH indices have low predictive ability. From April, when a warm season starts and mosquito activity intensifies, the correlation rapidly increases, reaching a maximum of –0.50 for VCI and 0.60 for TCI values during June to July (weeks 24–28). After these maximums, the correlation gradually decreases to a near-zero level by the beginning of the next cool season in November after week 40. In addition to a correct reflection of the timing and intensity of the malaria–VH relationship, the correlation with VCI and TCI values correctly explains the direction of mosquito reaction to satellite-based proxy. A negative correlation of *D* with VCI indicates that more malaria cases (*D* is above the trend) are developing for dryer condition (VCI $<$ 50 or reduced vegetation greenness, equation (2)). Conversely, fewer malaria cases (*D* is below the trend) are recorded for moist conditions (VCI $>$ 50 or larger vegetation greenness, equation (2)). This confirms that, in average wet climates, excessive rainfall during the monsoon season negatively affects mosquito activity and their ability to transmit malaria. Regarding thermal conditions, a larger number of malaria cases (*D* is above the trend) is associated with a TCI value greater than 50, which indicates cooler weather (equation (3)). A smaller number of malaria cases (*D* below the trend) is associated with lower TCI values (below 50, hotter weather, equation (3)) (Rahman *et al.* 2006, Salazar *et al.* 2007).

Correlation of *D* with VCI and TCI values for minor malaria divisions (Dhaka, Sylhet) showed that the PCC seasonal dynamics is erratic. However, it is important to emphasize that the PCC dynamic is similar in both divisions, and in some periods, the malaria–VH relationship is quite strong (PCC = –0.5, –0.6 for weeks 28 and 40). Therefore, the investigation of VCI and TCI values as predictors was performed for both major and minor malaria divisions.

Since Chittagong division contributes more than 90% of the malaria cases of the entire country cases, correlation of D with VCI and TCI values for the entire Bangladesh region shows a similar value to Chittagong correlation dynamics, especially for TCI values. The other two divisions (Dhaka and Sylhet), with erratic behaviour of correlation dynamics, also make a contribution to the entire Bangladesh region correlation of D with VCI and TCI values. As a result, the entire Bangladesh region correlation of D with TCI is shifted slightly to the later months (weeks 36–40), reflecting an elevated correlation (0.5–0.6) during this period in Dhaka and Sylhet divisions. The Bangladesh correlation between the D and VCI values also showed some shift in peak correlation to a later period and bi-modal dynamics.

Finally, it is important to emphasize, first, the dynamics of the PCCs in figure 5 correctly identify the seasonal cycle of malaria transmission in relation to the Bangladesh climate: increase of the absolute PCC values at the beginning of the warm season, maximum PCC during the critical period (middle of the malaria season) and decrease of PCC during the end of the warm season. Second, the maximum PCC of D with TCI is larger (0.8) then with VCI (–0.65), indicating that, in the wet climate of Bangladesh (annual precipitation 1000–3000 mm), temperature has a greater impact on the number of malaria cases than moisture. Third, lower TCI values (hotter weather, see equation (3)) during the critical period is an appropriate predictor of a smaller number of malaria cases (D below the trend) in Chittagong division and Bangladesh. Fourth, the D correlation with VCI is negative, indicating that dryer conditions (reduced greenness, see equation (2)) are conducive to more malaria cases. This is a very revealing fact, implying that in extremely wet climates, such as Bangladesh, abundant moisture levels hamper the mosquito's ability to transmit malaria.

6.3 *Regression analysis*

The results of correlation analysis (as shown in figure 5) were used to develop regression equations of D versus TCI and VCI values for each division and the entire Bangladesh region. Several options were investigated using either TCI (thermal condition) or VCI (moisture condition) values only, or both indices, for the weeks of the highest correlation (Rahman *et al.* 2006, Salazar *et al.* 2007). The general form of the regression equation when both indices were used is

$$Q = a_0 + a_1(\mathrm{TCI}_i) + a_2(\mathrm{VCI}_i), \tag{6}$$

where a_0, a_1 and a_2 are coefficients, i is the week number and Q is the predicted number of malaria cases (%) deviation from trend.

The tested variables are presented in table 3, with the corresponding multiple correlation coefficient (MCC), root mean square error (RMSE) and F criteria (to test statistical hypothesis). Analysis indicates that, for the two minor divisions (Dhaka and Sylhet), the MCC is not much different than for individual weeks, but the RMSE is quite large (30–35%). Such high errors could be expected since the area of the minor divisions is very remote; the population is not large and is spread over diversified ecosystems and environmental conditions. In spite of large RMSEs, several models were selected for further analysis. For Dhaka division, models 2 and 3 showed slightly higher MCC and lower RMSE values then others. Model 3 has some advantages in terms of early indication (week 28) of possible malaria epidemic. For the Sylhet division, model 3 was selected with the best estimates.

Table 3. Investigated variables, PCCs and MCCs of the models ($D = f$ (TCI, VCI)).

		Correlation coefficient of D with: TCI, for weeks						VCI, for weeks							
Division	Model	17	26	28	30	39	41	17	26	36	39	40	MCC	RMSE	F value
Chittagong	1		0.59										0.59	6.26	4.20
	2								0.48				0.48	6.78	2.40
	3		0.59						0.48				0.61	6.57	2.04
	4		0.59		0.42								0.63	6.45	2.25
Dhaka	1			0.50									0.53	33.77	3.07
	2						0.64						0.65	30.33	5.73
	3			0.50			0.64						0.65	29.09	3.96
	4			0.50								0.24	0.53	36.01	2.77
	5						0.64					0.24	0.66	31.74	1.37
Sylhet	1					0.50							0.50	33.02	2.69
	2					0.50					0.39		0.50	35.30	1.17
	3	0.29				0.50							0.68	29.86	3.03
Bangladesh	1						0.79						0.79	11.48	13.66
	2						0.79			0.66			0.80	12.10	6.25
	3									0.66			0.66	14.20	6.16
	4							0.60		0.66			0.69	14.71	3.10
	5		0.35				0.79						0.80	12.24	6.03

Table 4. Variables and statistical measures of the best models from table 3.

		Regression coefficient									
			t-statistics								
			Slope		Critical value		Variable				
Division	Model	a_0	a_1	a_2	5%	10%	1	2	MCC	RMSE	*F* value
Bangladesh	1	69.34	0.17		1.83	1.38	TCI_{41}		0.79	11.48	13.66
	5	67.36	0.20	3.12	1.83	1.38	TCI_{26}	TCI_{41}	0.80	12.24	6.03
Chittagong	3	96.50	1.24	0.53	1.83	1.38	TCI_{26}	VCI_{26}	0.61	6.57	2.04
	4	88.94	1.58	0.73	1.83	1.38	TCI_{26}	TCI_{30}	0.63	6.45	2.25
Dhaka	3	93.47	1.30	1.94	1.83	1.38	TCI_{28}	TCI_{41}	0.73	29.09	3.96
Sylhet	3	90.60	1.67	2.24	1.83	1.38	TCI_{17}	TCI_{39}	0.68	29.86	3.03

The smallest RMSEs were for the main malaria division of the country, Chittagong as well as for the entire Bangladesh region. For Chittagong division, the best models based on the MCC, RMSE and *F* parameters were 3 and 4. Model 4 (TCI_{26} and TCI_{30} predictors) provides slightly larger MCC and smaller RMSE values. But in terms of timeliness of prediction, model 3 provides advanced warning. For the entire Bangladesh region, three models showed similar results of estimates: model 1 (predictor TCI_{41}), 2 (TCI_{41} and VCI_{36}) and 5 (TCI_{26} and TCI_{41}). Although model 1 has the smallest RMSE, model 5 was selected for further analysis since one of the predictors (TCI for week 26) provides an early indication of a malaria epidemic.

Final equations of the best accepted models for the three malaria prone divisions and the entire Bangladesh region are shown in table 4. In addition to analysis of MCC, RMSE and *F* value, we also used the *t*-test for regression coefficients with a significance of 5% and 10%. Following table 4, both Bangladesh models have statistically significant *t*-test values for variable TCI_{41} (a_1 for model 1 and a_2 for model 5). Although a_1 in model 5 is not statistically significant, the TCI_{26} variable was included since it provides an early indication of a malaria epidemic. For Chittagong division, model 4 was selected as the best since the *t*-test value of 1.58 for the predictor TCI_{26} was higher than a critical value (1.38) with 10% significance. It is interesting to note that, when TCI and VCI values for the same week 26 were selected as predictors in model 3, the performance of the statistically significant predictor TCI_{26} deteriorated compared to model 4. For the Dhaka and Sylhet division models the second predictors (TCI_{41} and TCI_{39} correspondingly) showed statistical significance at 5% critical value.

7. Model validation

Further analysis included independent validation of models (table 4). Since the training data is short, the jackknife technique was used as a validation tool. For each model, one year of malaria and satellite data were excluded from the 1992–2001 dataset. A model ($Q = f$ (VCI, TCI)) was developed leaving one year out, and this model was applied to the removed year to predict the deviation of the number of malaria cases from trend (Q) based on satellite data of the eliminated year. Then, the eliminated year was returned to the dataset and the next year was removed for model development and testing (Bruce 1987). Each year's data were removed one at a time, and the candidate model was fitted nine times to the eliminated year. As the result of this procedure, nine independent predictions were obtained.

Finally, in each of the predictions, the number of malaria cases (*P*) for the eliminated year (*i*) was estimated from equation (7). In addition, the coefficient of determination (R^2) for the number of independently predicted and observed malaria cases, the bias (*B*), percentage of relative bias (RB) and root mean square error (RMSE) for this year was estimated using equations (8), (9) and (10):

$$P_i = Y_{\text{trend}}(Q_i/100), \tag{7}$$

$$B_i = P_i - Y_i, \tag{8}$$

$$\text{RB}_i = (B_i - \bar{B})100/Y_i \tag{9}$$

and

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{10}(B_i)^2}{10}}, \tag{10}$$

where *P* is the predicted malaria cases (%); $\bar{B}$ is the average bias for all years; and the RMSE is a measure of the precision of the predicted value and should be as small as possible for unbiased precise prediction.

Models were independently tested: 1 and 5 for Bangladesh, 3 and 4 for Chittagong and 3 for Dhaka and Sylhet division (table 4). In the estimation of model performance, we followed the rules: (1) for the entire model, R^2 greater than 70 and the RMSE was less than 15% and (2) for individual years, the bias (*B*) was less than 2% and the relative bias (RB) was less than 10%.

The independently validated results presented in table 5 and figure 6 show that, for the entire models, the R^2 and RMSE criteria have been met for Bangladesh, Chittagong (model 3) and Dhaka. However, the analysis of the annual model's performance showed that only Bangladesh and Chittagong models perform reliably. In 8 years for the Chittagong division and 9 years (out of 10 tested) for Bangladesh, the bias was less than 2% and the relative bias was less than 10%. Regarding models for minor malaria regions, the annual results of testing are negative since, as seen in table 5, the relative bias indicates a strong deterioration of the model's performance after 1996 (Sylhet RB was 8–16% and Dhaka RB was 19–33%), while prior to 1996, the relative bias was 50% less. Such an explanation can be given to the fact that the Bangladesh Government developed very comprehensive measures to combat malaria in Sylhet and Dhaka. During 1996–2001, this resulted in a considerable reduction of malaria cases (as shown in figure 4), and the data must be interpreted with caution because of the decline in surveillance activities in the country over the past few years (WHO 1999).

The Bangladesh Government has also undertaken malaria-fight measures in 13 districts from Chittagong, Sylhet and Dhaka divisions. However, the effectiveness was not as good as in the minor malaria regions because large numbers of people were exposed to malaria, especially among the poor in Chittagong division (93% of the entire Bangladesh region people affected). The lower effectiveness of the malaria combat measures is also indicated by an increase in the number of malaria cases during 1992–2001 (as shown in figure 4).

Table 5. Independent evaluation of the best regression model using the jackknife technique.

(a) Chittagong: $R^2 = 0.71$ (between simulated and observed malaria cases), RMSE = 1.20.

Year excluded	Malaria observed (Y)	Cases predicted (P) (%)	Bias ($B = Y - P$) (%)	Relative Bias (RB) (%)	Estimated trend (Y_{trend}) (%)	Predicted deviation from trend (Q) (%)
1992	18.18	17.49	–0.69	–2.67	19.20	91.10
1993	18.60	18.44	–0.16	0.25	19.76	93.33
1994	21.91	23.51	1.60	8.23	20.31	115.75
1995	21.95	19.06	–2.89	–12.22	20.86	91.39
1996	22.60	19.22	–3.38	–14.05	21.42	89.73
1997	18.94	19.93	0.99	6.28	21.97	90.70
1998	24.23	24.09	–0.14	0.25	22.52	106.95
1999	24.23	25.03	0.80	4.15	23.07	108.50
2000	23.14	25.33	2.19	10.33	23.63	107.18
2001	23.14	22.79	–0.35	–0.65	24.18	94.24
Mean	21.69	21.49	–0.20	–0.01	21.69	98.89

(b) Bangladesh: $R^2 = 0.88$ (between simulated and observed malaria cases), RMSE = 1.23.

Year excluded	Malaria observed (Y)	Cases predicted (P) (%)	Bias (B) (%)	Relative Bias (RB) (%)	Estimated trend (Y_{trend}) (%)	Predicted deviation from trend (Q) (%)
1992	6.02	5.71	–0.31	–6.39	6.24	91.58
1993	7.60	7.58	–0.02	–1.26	7.33	103.45
1994	10.24	10.26	0.02	–0.53	8.41	122.05
1995	10.45	9.61	–0.84	–8.84	9.50	101.11
1996	8.70	8.19	–0.51	–6.75	10.59	77.35
1997	7.17	9.83	2.66	36.06	11.67	84.27
1998	13.70	11.82	–1.88	–14.33	12.76	92.61
1999	16.50	15.86	–0.64	–4.38	13.85	114.49
2000	15.53	17.15	1.62	9.91	14.93	114.85
2001	15.39	16.07	0.68	3.93	16.02	100.34
Mean	11.13	11.21	0.08	0.74	11.13	100.21

(c) Sylhet: $R^2 = 0.37$ (between simulated and observed malaria cases), RMSE = 4.30.

Year excluded	Malaria observed (Y)	Cases predicted (P) (%)	Bias (B) (%)	Relative bias (RB) (%)	Estimated trend (Y_{trend}) (%)	Predicted deviation from trend (Q) (%)
1992	9.60	15.50	–5.90	9.53	15.91	97.41
1993	12.76	12.50	0.26	7.17	14.88	83.98
1994	21.25	16.37	4.88	4.30	13.85	118.21
1995	20.67	13.67	7.00	4.42	12.82	106.66
1996	12.12	7.49	4.63	7.55	11.78	63.56
1997	7.40	9.68	–2.28	12.36	10.75	90.03
1998	7.03	6.74	0.29	13.01	9.72	69.34
1999	5.85	12.36	–6.51	15.63	8.69	142.27
2000	8.00	6.85	1.15	11.43	7.65	89.59
2001	8.00	6.98	1.02	11.43	6.62	105.50
Mean	11.27	10.81	0.45	9.68	11.27	96.66

(*Continued*)

Table 5. (*Continued.*)

(d) Dhaka: $R^2 = 0.83$ (between simulated and observed malaria cases), RMSE = 1.09.

Year excluded	Malaria observed (Y)	Cases predicted (P) (%)	Bias (B) (%)	Relative bias (RB) (%)	Estimated trend (Y_{trend}) (%)	Predicted deviation from trend (Q) (%)
1992	5.40	6.92	–1.52	10.53	7.24	95.57
1993	7.60	7.64	–0.04	7.48	6.54	116.80
1994	7.70	8.53	–0.83	7.39	5.84	146.13
1995	6.90	5.46	1.44	8.24	5.14	106.31
1996	3.00	1.37	1.63	18.96	4.44	30.87
1997	2.03	3.50	–1.47	28.01	3.74	93.45
1998	2.50	2.56	–0.06	22.75	3.05	83.97
1999	2.40	3.56	–1.16	23.70	2.35	151.59
2000	1.70	2.16	–0.46	33.45	1.65	130.71
2001	1.70	1.07	0.63	33.45	0.95	112.43
Mean	4.09	4.28	–0.18	19.40	4.09	106.78

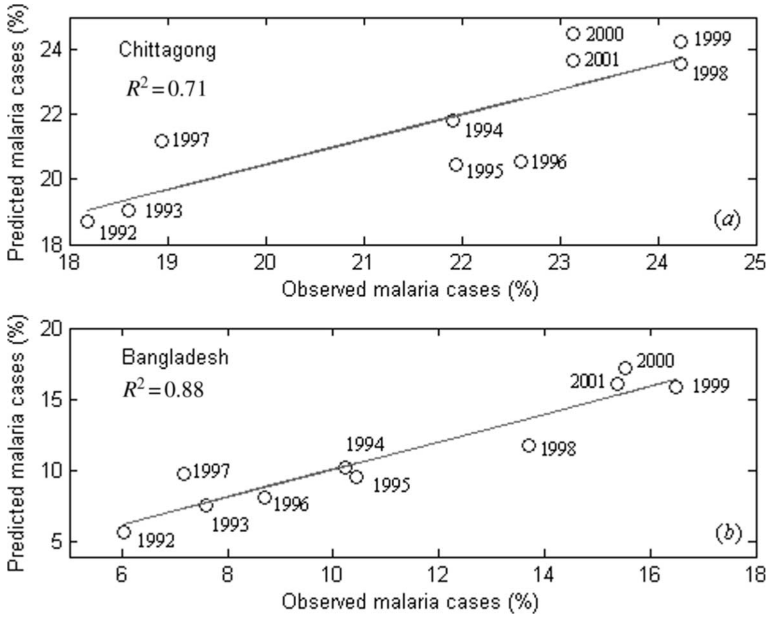

Figure 6. Correlation between independently predicted and observed number of malaria cases (% of malaria cases from the total number of people who come to the regional hospitals with fever).

It would be important to emphasize that, although Chittagong and Bangladesh models met both performance criteria (for the entire model and individual years), in some years, the bias and relative bias exceed the threshold's level. Further analysis is focused on these years: 1997 (RB = 11%) and 1997 (RB = 36%) for Chittagong, 1998 (RB = 14%) for Bangladesh. In 1997, models for both regions overestimated the number of malaria-affected people and in 1998 this number was underestimated in

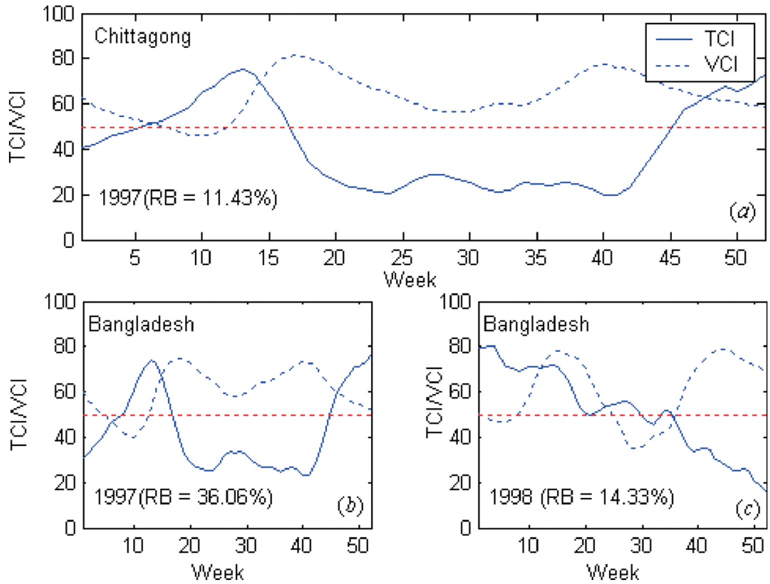

Figure 7. Dynamics of TCI during the years of RB exceeding the evaluation threshold.

Bangladesh. First, we should emphasize that 1997/98 was a strong El Niño year (positive sea-surface temperature anomaly in the tropical Pacific) (Barnston *et al.* 1997, Kovats *et al.* 2003). As a result, the southeastern monsoon in Bangladesh was delayed by one month, resulting in a long period of extremely hot and dry weather. Figure 7 shows TCI and VCI dynamics in Chittagong and Bangladesh. During the investigated years, severe drought (TCI = 15–30) developed from June through to October. Such conditions produced a much stronger impact on mosquito development, disrupting reproductive mosquito cycles and the intensity of malaria transmission, which resulted in a smaller number of malaria cases compared to the prediction. To characterize such extreme conditions, the model's predictors should characterize longer periods (several weeks and even months), but this is not possible now due to a limited statistical sample. Unfortunately, we did not find a reasonable explanation for the 1998 smaller number of predicted malaria cases in the entire Bangladesh region. It is important to note that there is some contradiction between Chittagong and Bangladesh predictions in 1998: the Chittagong (contributes 97% of malaria cases) predictor met the selected criteria (RB = 3.14%) and the Bangladesh prediction did not (RB = 14.33%).

8. Conclusions

AVHRR-based VH indices characterizing moisture (VCI) and thermal (TCI) conditions were used during 1992–2001 as predictors for estimation of malaria cases in the three administrative divisions and compared with the entire Bangladesh region. Correlation between the number of malaria case deviations from trend (*D*) in Bangladesh and the Chittagong division (the most affected with malaria) with TCI and VCI was similar: strong during June through to October (main malaria season) and weak during November through to May. The transition period for malaria is from April (warm season starts), when the correlation rapidly increases, through to June (weeks 24–26), when the correlation reaches a maximum. After the maximum, the correlation gradually decreases, reaching nearly zero in November (after week 40). It was found that the number of malaria cases was more sensitive to thermal (TCI) then moisture (VCI) conditions. Therefore, statistical models were developed for prediction of the number of malaria cases based on TCI parameters. In minor divisions (Sylhet and Dhaka) where the population is distributed over a large area and malaria-fight measures provide positive results, the correlation dynamics were not well determined.

Correlation and regression analysis shows that a malaria epidemic can be carried out well ahead of an epidemic occurrence. These models can be used to predict malaria in districts and the entire Bangladesh region.

Further investigation might include high-resolution satellite data (sea-surface temperature (SST) and soil moisture) from the National Aeronautics and Space Administration's (NASA's) Moderate Resolution Imaging Spectroradiometer (MODIS), and Radar-Satellite (RADARSAT), investigation of such factors as sea-surface temperature and combing satellite data and weather data. Satellite technology is currently not available in Bangladesh. Therefore, VH index data delivered in real time to http://orbit.nesdis.noaa.gov/smcd/emcb can be used, together with the obtained equations for early detection of conditions that are suitable for malaria epidemics.

Acknowledgements

This study is supported by National Environmental Satellite Data and Information Service (NESDIS) award no. NA07NES4280009. The statements contained in this article are not the opinions of the funding agency or the USA government, but reflect the opinions of the authors.

References

Allard, R., 1998, Use of time-series analysis in infectious disease surveillance. *Bulletin of World Health Organization*, **76**, pp. 327–333.

Barnston, A.G., Chelliah, M. and Goldenberg, S.B., 1997, Documentation of a highly ENSO-related SST region in the equatorial Pacific. *Atmosphere – Ocean*, **3**, pp. 367–383.

Boëte, C. and Koella, J., 2002, A theoretical approach to predicting the success of genetic manipulation of malaria mosquitoes in malaria control. *Malaria Journal*, **1**, p. 3.

Bouma, M.J., 2003, Methodological problems and amendments to demonstrate effects of temperature on the epidemiology of malaria. A new perspective on the highland epidemics in Madagascar1972–1989. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **97**, pp.133–139.

Bruce, M.D., 1987, Uses and abuses of cross-validation in geostatistics. *Mathematical Geology*, **19**, pp. 241–248.

Byron, L.W., Louisa, R.B., Robert, K.W., Susan, M.P. and Paul, D.S., 1991, Spectral and spatial characterization of rice field mosquito habitat. *International Journal of Remote Sensing*, **12**, pp. 621–626.

Ceccato, P., Connor, S.J., Jeanne, I. and Thomson, M.C., 2005, Application of geographical information system and remote sensing in malaria risk. *Parasitologia*, **47**, pp. 81–96.

CHANDRAMOHAN, D., JAFFAR, S. and GREENWOOD, B., 2002, Use of clinical algorithms for diagnosing malaria. *Tropical Medicine and International Health*, **7**, pp. 45 –52.

CHILUNDO, B., SUNDBY, J. and AANESTAD, M., 2004, Analysing the quality of routine malaria data in Mozambique. *Malaria Journal*, **3**, p. 3.

CONNOR, S., THOMSON, M. and MOLYNEUX, D., 1999, Forecasting and prevention of epidemic malaria: new perspectives on an old problem. *Parasitologia*, **41**, pp. 439–448.

CRAIG, M.H., KLEINSCHMIDT, I., LE SUEUR, D. and SHARP, B,L., 2004, Exploring 30 years of malaria case data in KwaZulu-Natal, South Africa: part II. The impact of non-climatic factors. *Tropical Medicine and International Health*, **9**, pp.1258–1266.

ELIAS, M. and RAHMAN, M., 1987, The ecology of malaria carrying mosquito Anopheles Philippinensis Ludlow and its relation to malaria in Bangladesh. *Medical Research Council Bulletin*, *Bangladesh*, **13**, pp. 15–28.

FAIZ, M.A., YUNUS, E.B., RAHMAN, M.R., HOSSAIN, M.A., PANG, L.W., RAHMAN, M.E. and BHUIYA, S.N., 2002, Failure of national guidelines to diagnose uncomplicated malaria in Bangladesh. *American Journal of Tropical Medicine and Hygiene*, **67**, pp. 396–399.

GITHEKO, A., LINDSAY, S., CONFALONIERI, U. and PATZ, J., 2000, Climate change and vector-borne diseases: a regional analysis. *Bulletin of World Health Organization*, **78**, pp. 200–207.

GRAMICCIA, G., 1952, Final comprehensive report, Pakistan, E. Bengal malaria control demonstration team, Mymensing district. *Pakistan Journal of Health*, **2**, pp. 61–88.

HAY, S.I., ROGERS, D.J., SHANKS, G.D., MYERS, M.F. and SNOW, R.W., 2001, Malaria early warning in Kenya. *Trends in Parasitology*, **17**, pp. 95–99.

HAY, S.I., ROGERS, D.J., RANDOLPH, S.E., STERN, D.I, COX, J., SHANKS, G.D. and SNOW, R.W., 2002, Hot topic or hot air? Climate change and malaria resurgence in east African highlands. *Trends in Parasitology*, **18**, pp. 530–534.

INGRID, V.F. and VAN, D.B., 2004, Drug resistance in plasmodium falciparum from the Chittagong Hill Tracts, Bangladesh. *Tropical Medicine and International Health*, **9**, pp. 680–687.

KAYA, S., PULTZ, T.J., MBOGO, C.M., BEIER, J.C. and MUSHINZIMANA, E., 2002, The use of radar remote sensing for identifying environmental factors associated with malaria risk in coastal Kenya. In *IGARSS*, June 2002, pp. 24–28.

KIDWELL, K.B., 1997, *Global Vegetation Index Users Guide NOAA*. Technical Report, NOAA, US Department of Commerce, Suitland, Maryland 65.

KOGAN, F., 2001, Operational space technology for global vegetation assessment. *Bulletin of the American Meteorological Society*, **82**, pp. 1949–1964.

KOGAN, F., 2002, World droughts in the new millennium from AVHRR-based vegetation health indices. *Eos*, **83**, pp. 557–564.

KOVATS, R.S., BOUMA, M.J, HAJAT, S., WORRALL, E. and HAINES, A., 2003, El Nino and health. *Lancet*, **362**, pp. 1481–1489.

MCMICHAEL, A.J., HAINES, A. and SLOOFF, R., 1996, Climate change and human health. *World Health Organization*, *Geneva*, *Switzerland*, **29**.

MONTANARI, R., BANGALI, M., TALUKDER, K., BAQUI, A., MASHEWARY, N., GOSH, A., RAHMAN, M. and MAHMOOD, A., 2001, Three case definitions of malaria and their effect on diagnosis, treatment and surveillance in Cox's Bazar district, Bangladesh. *Bulletin of the World Health Organization*, **79**, pp. 648–656.

NAGPAL, B. and SHARMA, V., 1995, *Indian Anophelines*, pp. 416–423 (New Delhi: Oxford and IBH Publishing).

NAJERA, J.A., KOUZNETZSOV, R.L. and DELACOLLETTE, C., 1998, *Malaria Epidemics: Detection and Control, Forecasting and Prevention*. World Health Organization, Geneva, Switzerland, WHO/MAL/98.

PAMPANA, E. (Ed.), 1969, *A Text Book of Malaria Eradication*, pp. 17–63 (London, UK: Oxford University Press).

Paresul, A., 2008, *Malaria Country Report*. Malaria and Parasitic Disease Control Unit, Directorate General of Health Services, Bangladesh.

Rahman, A., Kogan, F. and Roytman, L., 2006, Analysis of malaria cases in Bangladesh with remote sensing data. *American Journal of Tropical Medicine and Hygiene*, **74**, pp. 17–19.

Remme, J.H.F., Binka, F. and Nabarro, D., 2001, Toward a framework and indicators for monitoring Roll Back Malaria. *American Journal of Tropical Medicine and Hygiene*, **64**, pp. 76–84.

Rogers, D.J., Randolph, S.E., Snow, R.W. and Hay, S.I., 2002, Satellite imagery in the study and forecast of malaria. *Nature*, **41**, pp. 710–715.

Rosenberg, R. and Maheswary, N., 1982, Forest malaria in Bangladesh, I. Parasitology. *American Journal of Tropical Medicine and Hygiene*, **31**, (a) pp. 175–182, (b) pp. 183–191, (c) pp. 192–201.

Salazar, L., Kogan, F. and Roytman, L., 2007, Using vegetation health indices and partial least squares method for estimation of corn yield. *International Journal of Remote Sensing*, **29**, pp. 175–189.

Salazar, L., Kogan, F. and Roytman, L., 2008, Use of remote sensing data for estimation of winter wheat yield in the United States. *International Journal of Remote Sensing*, **28**, pp. 3795–3811.

Smith, D. and Mckenzie, E., 2004, Statics and dynamics of malaria infection in Anopheles mosquitoes. *Malaria Journal*, **3**, p. 13.

Thomson, M.C. and Connor, S.J., 2001, The development of malaria early warning systems for Africa. *Trends in Parasitology*, **17**, pp. 438–445.

Thomson, M.C., Connor, S.J., O'Niell, K. and Meert, J.P., 2000, Environmental information for epidemic prediction. *Parasitology Today*, **16**, pp. 137–138.

Wickramasinghe, A.R., Gunawardena, D.M. and Mahawithanage, S.T., 2002, Use of routinely collected past surveillance data in identifying and mapping high risk areas in a malaria endemic area of Sri Lanka. *Southeast Asian Journal of Tropical Medicine and Public Health*, **33**, pp. 678–684.

World Health Organization (WHO), 1999, *The World Health Report, ROLLING BACK MALARIA* (Geneva, Switzerland: WHO). Available online at: http://www.rbm.who.int/docs/whr99.htm (accessed 18 March 2009).

World Health Organization (WHO), 2002, *Final Report on the Third Meeting of the RBM Technical Resource Network on Epidemic Prevention and Control* (Geneva, Switzerland: WHO).

Zhou, G., Minakawa, N., Githeko, A.K. and Yan, G., 2004, Association between climate variability and malaria epidemics in the east African highlands. *Proceedings of the National Academy of Sciences USA*, **101**, pp. 2375–2380.